《概率论与数理统计作业》

姓名：尹超 学号：2023K8009926003 专业：人工智能 班级：2313

目录

§ 1 12.02

题目描述

Show that in simple linear regression,

$$\text{Coefficient of Determination} = (\text{Coefficient of correlation})^2$$

Where Coefficient of Determination is defined as:

$$\text{Coefficient of Determination} = \frac{SS_{\text{total}} - SS_{\text{err}}}{SS_{\text{total}}}$$

证明在简单线性回归中，

$$\text{决定系数} = (\text{相关系数})^2$$

其中决定系数定义为：

$$\text{决定系数} = \frac{SS_{\text{total}} - SS_{\text{err}}}{SS_{\text{total}}}$$

解答过程

在简单线性回归中，我们将因变量 y 与自变量 x 之间的关系建模为：

$$y = \beta_0 + \beta_1 x + \epsilon$$

其中：- β_0 是截距。- β_1 是斜率。- ϵ 是误差项。

决定系数（R²）定义为：

$$R^2 = \frac{SS_{\text{total}} - SS_{\text{err}}}{SS_{\text{total}}}$$

其中：- SS_{total} 是 y 的总平方和，表示数据的总方差。- SS_{err} 是残差平方和，表示回归模型无法解释的方差。

相关系数（r）是皮尔逊相关系数，用来衡量 x 和 y 之间线性关系的强度。在简单线性回归中，相关系数 r 的公式为：

$$r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

其中：- $\text{Cov}(x, y)$ 是 x 和 y 的协方差。- σ_x 和 σ_y 分别是 x 和 y 的标准差。

决定系数（R²）可以写成：

$$R^2 = \frac{SS_{\text{reg}}}{SS_{\text{total}}}$$

其中回归平方和 SS_{reg} 是总平方和和误差平方和的差：

$$SS_{\text{reg}} = SS_{\text{total}} - SS_{\text{err}} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

因此，决定系数 R^2 与相关系数 r 的关系是：

$$r^2 = \frac{SS_{\text{reg}}}{SS_{\text{total}}}$$

所以我们可以得出结论：

$$R^2 = r^2$$

这证明了在简单线性回归中，决定系数（R^2）等于相关系数（r）的平方。

题目描述

假设 $Y_i = b_0 + b_1 x_i + e_i$，其中 $e_1, e_2, \ldots, e_n$ 独立同分布，且 $e_i \sim N(0, \sigma^2)$。
找到 b_0 和 b_1 的最大似然估计并验证它们是最小二乘估计。

解答过程

首先，写出似然函数：

$$L(b_0, b_1, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - b_0 - b_1 x_i)^2}{2\sigma^2}\right)$$

取对数得到对数似然函数：

$$\ell(b_0, b_1, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - b_0 - b_1 x_i)^2$$

对 b_0 和 b_1 求偏导数并令其等于零：

$$\frac{\partial \ell}{\partial b_0} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - b_0 - b_1 x_i) = 0$$

$$\frac{\partial \ell}{\partial b_1} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - b_0 - b_1 x_i)x_i = 0$$

解这两个方程可以得到 b_0 和 b_1 的最大似然估计：

$$b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$b_0 = \bar{Y} - b_1 \bar{x}$$

这正是最小二乘估计的公式，因此最大似然估计和最小二乘估计是一致的。

§ 2 12.04

题目描述

在简单线性回归模型中，证明 ANOVA 中的 F 统计量是用于检验 $H_0 : b_1 = 0$ 与 $H_1 : b_1 \neq 0$ 的 T 统计量的平方。

证明

在简单线性回归中，模型为：

$$Y_i = b_0 + b_1 x_i + e_i$$

假设检验 $H_0 : b_1 = 0$ 与 $H_1 : b_1 \neq 0$，T 统计量定义为：

$$t = \frac{b_1}{\text{SE}(b_1)}$$

其中，$\text{SE}(b_1)$ 是 b_1 的标准误。

ANOVA 中的 F 统计量定义为：

$$F = \frac{\text{MS}_{\text{reg}}}{\text{MS}_{\text{err}}}$$

其中，MS_{reg} 是回归均方，MS_{err} 是误差均方。

在简单线性回归中，MS_{reg} 和 MS_{err} 可以表示为：

$$\text{MS}_{\text{reg}} = \frac{SS_{\text{reg}}}{1}$$

$$\text{MS}_{\text{err}} = \frac{SS_{\text{err}}}{n-2}$$

其中，SS_{reg} 是回归平方和，SS_{err} 是误差平方和。

由于 $SS_{\text{reg}} = b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2$，我们有：

$$F = \frac{b_1^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\frac{SS_{\text{err}}}{n-2}}$$

注意到 $\text{SE}(b_1) = \sqrt{\frac{SS_{\text{err}}}{(n-2)\sum_{i=1}^n (x_i - \bar{x})^2}}$，我们可以得到：

$$t^2 = \left(\frac{b_1}{\text{SE}(b_1)}\right)^2 = \frac{b_1^2}{\frac{SS_{\text{err}}}{(n-2)\sum_{i=1}^n (x_i - \bar{x})^2}} = F$$

因此，ANOVA 中的 F 统计量是用于检验 $H_0 : b_1 = 0$ 与 $H_1 : b_1 \neq 0$ 的 T 统计量的平方。

§ 3 12.09

题目描述

在 Deming 回归中，当 $\delta = 1$ 时，请证明残差平方和

$$RSS = \frac{1}{1+\beta_1^2}\sum_{i=1}^{n}(y_i - \beta_1 x_i - \beta_0)^2$$

注意：除了笔记中的方法，也可以用直角三角形面积方法。

证明

在 Deming 回归中，残差平方和 (RSS) 的定义是：

$$RSS = \sum_{i=1}^{n}\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{1+\beta_1^2}$$

当 $\delta = 1$ 时，Deming 回归的目标是最小化以下目标函数：

$$Q = \sum_{i=1}^{n}\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{1+\beta_1^2}$$

我们可以将其重写为：

$$Q = \frac{1}{1+\beta_1^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

因此，残差平方和 (RSS) 可以表示为：

$$RSS = \frac{1}{1+\beta_1^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

这证明了在 $\delta = 1$ 的情况下，Deming 回归的残差平方和的公式。

题目描述

设 ρ 和 r 分别为相关系数和皮尔逊相关系数。使用 Fisher 变换，我们已经得到了 $z(\rho) = \frac{1}{2}\log_e \frac{1+\rho}{1-\rho}$ 的渐近 $1-\alpha$ 置信区间为 $[\zeta_{\text{lower}}, \zeta_{\text{upper}}]$，其中

$$\zeta_{\text{lower}} = z(r) - z_{(1-\alpha/2)}\sqrt{n-\frac{1}{3}}, \quad \zeta_{\text{upper}} = z(r) + z_{(1-\alpha/2)}\sqrt{n-\frac{1}{3}}$$

请进一步证明 ρ 的渐近置信区间为：

$$\left[\frac{\exp\{2\zeta_{\text{lower}}\}-1}{\exp\{2\zeta_{\text{lower}}\}+1}, \frac{\exp\{2\zeta_{\text{upper}}\}-1}{\exp\{2\zeta_{\text{upper}}\}+1}\right]$$

证明

根据 Fisher 变换，$z(\rho) = \frac{1}{2}\log_e \frac{1+\rho}{1-\rho}$。我们已经得到了 $z(\rho)$ 的置信区间 $[\zeta_{\text{lower}}, \zeta_{\text{upper}}]$。
为了得到 ρ 的置信区间，我们需要对 $z(\rho)$ 的置信区间进行反变换。即：

$$\rho = \frac{\exp\{2z(\rho)\}-1}{\exp\{2z(\rho)\}+1}$$

因此，ρ 的置信区间为：

$$\left[\frac{\exp\{2\zeta_{\text{lower}}\}-1}{\exp\{2\zeta_{\text{lower}}\}+1}, \frac{\exp\{2\zeta_{\text{upper}}\}-1}{\exp\{2\zeta_{\text{upper}}\}+1}\right]$$

§4 12.11

题目描述

假设 X_1 和 X_2 是方差为 σ^2 的不相关随机变量，使用矩阵方法证明 $Y = X_1 + X_2$ 和 $Z = X_1 - X_2$ 是不相关的。

证明

设 $\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$，则 $\mathbf{Y} = \begin{pmatrix} Y \\ Z \end{pmatrix}$ 可以表示为：

$$\mathbf{Y} = \mathbf{A}\mathbf{X}$$

其中，$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$。

首先，计算 $\mathbf{X}$ 的协方差矩阵 $\Sigma_{\mathbf{X}}$：

$$\Sigma_{\mathbf{X}} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

然后，计算 $\mathbf{Y}$ 的协方差矩阵 $\Sigma_{\mathbf{Y}}$：

$$\Sigma_{\mathbf{Y}} = \mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}^T$$

计算 $\mathbf{A}\Sigma_{\mathbf{X}}$：

$$\mathbf{A}\Sigma_{\mathbf{X}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} = \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & -\sigma^2 \end{pmatrix}$$

然后计算 $\mathbf{A}\Sigma_{\mathbf{X}}\mathbf{A}^T$：

$$\Sigma_{\mathbf{Y}} = \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & -\sigma^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2\sigma^2 & 0 \\ 0 & 2\sigma^2 \end{pmatrix}$$

因此，$\Sigma_{\mathbf{Y}}$ 是对角矩阵，这表明 Y 和 Z 是不相关的。

题目描述

考虑拟合曲线 $y = b_0x + b_1x^2$ 到点 (x_i, y_i)，其中 $i = 1, \ldots, n$。

1. 使用矩阵形式找到 b_0 和 b_1 的最小二乘估计的表达式。2. 拟合值是什么？

解答

1. 使用矩阵形式找到 b_0 和 b_1 的最小二乘估计的表达式。

设 $\mathbf{y}$ 是观测值向量，$\mathbf{X}$ 是设计矩阵，$\mathbf{b}$ 是参数向量，$\mathbf{e}$ 是误差向量，则有：

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

其中，

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_1 & x_1^2 \\ x_2 & x_2^2 \\ \vdots & \vdots \\ x_n & x_n^2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

最小二乘估计 $\mathbf{b}$ 满足：

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

2. 拟合值是什么？

拟合值 $\hat{\mathbf{y}}$ 为：

$$\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$